Self-averaging property of queuing systems

Alexander Rybko, Senya Shlosman, Alexander Vladimirov

Self-averaging property of queuing systems.

Alexandre Rybko
Institute for the Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia
rybko@iitp.ru

Senya Shlosman
Institute for the Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia,
shlos@iitp.ru
Centre de Physique Theorique
– Unite Mixte de Recherche (UMR 6207) du CNRS
et des Universites Aix-Marseille et du Sud Toulon-Var,
shlosman@cpt.univ-mrs.fr

Alexandre Vladimirov
Institute for the Information Transmission Problems,
Russian Academy of Sciences, Moscow, Russia
vladim@iitp.ru

April 26, 2006

Abstract

We establish the averaging property for a queuing process with one server, $M(t)/G/1$. It is a new relation between the output flow rate and the input flow rate, crucial in the study of the Poisson Hypothesis. Its implications include the statement that the output flow always possesses more regularity than the input flow.

Keywords: service time, stochastic kernel, non-linear equation, self-averaging.

1 Introduction

The Poisson Hypothesis deals with large queuing systems. It is the statement that for certain large networks the input flow to any given node is approximately Poissonian with constant rate.

In the paper [RS1] the Poisson Hypothesis is proven for some simple queuing networks. One of the main technical ingredients of this proof is the following non-linear averaging relation for $M(t)/GI/1$ queuing process with one server:

$$b(t) = [\lambda(\cdot) * q_{\lambda,t}(\cdot)](t). \tag{1}$$

Here $*$ stays for convolution: for two functions $a(\cdot), b(\cdot)$ it is defined as

$$[a(\cdot) * b(\cdot)](t) = \int a(t-x)\, b(x)\, dx.$$

In order to explain the rest of the relation (1) we now introduce notation for our server. Here $\lambda(t)$, $-\infty < t < \infty$ is the rate of the Poisson process of moments of arrivals of customers to our server. If the server is busy, the customer waits in line for his turn; the service discipline is First-In-First-Out (FIFO). The service time η_i for the i-th customer is supposed to be random, while the random variables η_i are independent identically distributed random variables, with common distribution η. Upon completion of the service the customer exits the system. The exit flow is of course not Poissonian in general; yet its rate $b(t)$ is defined. The claim (1) is that the functions λ and b are related via the convolution with the kernel $q_{\lambda,t}(x)$, which is supported by positive semiaxis:

$$q_{\lambda,t}(x) = 0 \text{ for } x < 0,$$

and, what is of crucial importance, is stochastic: for every t

$$\int q_{\lambda,t}(x)\, dx = 1. \tag{2}$$

The function $q_{\lambda,t}(x)$ depends on the function $\lambda(y)$ only via its restriction to the semiline $\{y \le t\}$. (Of course, it depends also on the law of η.) If the system is not overloaded, the family $q_{\lambda,t}(\cdot)$ of probability measures is also compact: for every $\varepsilon > 0$ there exists a threshold K such that $\int_0^K q_{\lambda,t}(x)\, dx >$

$1-\varepsilon$, uniformly in t. As it is explained in [RS1], the averaging relation (1) with *stochastic* kernel $q_{\lambda,t}$ does not hold in general for other disciplines. Some other examples of self-averaging violation are presented in Sect. 4. The importance of (1) lies in the fact that it implies the rate b is in a sense "smoother" than λ; in particular, it implies that $\sup b \leq \sup \lambda$ and $\inf b \geq \inf \lambda$, and it "almost" implies that the last inequalities are in fact strict.

The proof of $(1)-(2)$, given in [RS1], is quite complicated, being based on the validity of a certain combinatorial identity dealing with rod placements on the real line (see also [RS2] and [A]). The independence of the service times η_i is very important in this proof. The purpose of the present paper is to extend the relations $(1)-(2)$ to the case when the sequence η_i is not necessarily independent, but has a weaker property of being a stationary ergodic process, i.e. $M(t)/G/1$ queuing system. We replace the combinatorial identity by a stochastic one. In a sense, both in [RS1] and here we are making use of Fubini theorem, and we use here a different choice of coordinates, which enables our extension to the dependent case. The applications of the present result to the Poisson Hypothesis will be presented elsewhere.

It is noteworthy to repeat that the result of the present paper is again based on an identity, this time a stochastic one. To formulate it, consider the random variable V, which is a functional of the service process trajectory, ω, and which is defined as follows:

- if the realization ω is such that the server is idle at the moment $t=0$, then $V(\omega)=0$;
- if the server is occupied at $t=0$, then let us introduce

  $\hat{\eta}(\omega)$ to be the total service time required by the customer, who is being served at this moment $t=0$,

  $\hat{t}(\omega)<0$ to be the moment of the beginning of the occupation period of the server, which period contains the moment $t=0$,

  and finally put
  $$V(\omega)=\frac{1}{\lambda\left(\hat{t}(\omega)\right)\hat{\eta}(\omega)}.$$

We claim now that
$$\mathbb{E}\left(V(\omega)\right)\equiv 1, \tag{3}$$

provided only that the server is not overloaded, plus some general technical conditions. Being very general, the identity (3) is surprisingly non-evident, similarly to the rod placement identity!

2 The main result

2.1 Notation

Let $\mathcal{P}$ be a Poisson point process on $\mathbb{R}^1$ of arrival moments of the customers. It is a probability measure on the set $\Omega' = \{... < z_{-1} < z_0 < z_1 < ...\}$ of double-infinite sequences $z_i \in \mathbb{R}^1$, which are locally finite subsets of $\mathbb{R}^1$. Every such Poisson process $\mathcal{P}$ is defined by the choice of a measure dm on $\mathbb{R}^1$, $\mathcal{P} = \mathcal{P}_m$, and we will suppose that

$$dm = \lambda(x)\, dx,$$

where $\lambda > 0$ is the rate of the process $\mathcal{P}_m = \mathcal{P}_\lambda$.

Strictly speaking, the Poisson process is a measure on locally finite subsets $\phi \subset \mathbb{R}^1$. We consider it as a measure on sequences by defining the point z_0 to be the smallest positive point in ϕ.

We further assume that the process $\mathcal{T}$ of *positive* service times $\{\dots, \eta_{-1}, \eta_0, \eta_1, \dots\}$, independent of $\mathcal{P}_\lambda$, is given. We will assume that $\mathcal{T}$ is stationary and ergodic.

The total process we thus are interested in, is the process $\mathcal{S} = \mathcal{P}_\lambda \times \mathcal{T}$, which is a probability measure on the set $\Omega = \{\omega = \dots, (z_{-1}, \eta_{-1}), (z_0, \eta_0), (z_1, \eta_1), \dots\}$.

2.2 Nonlinear shift

In the special case when the rate λ equals to a constant ℓ, the process $\mathcal{S}$ is ergodic with respect to the shift transformation T_t on Ω :

$$\begin{aligned} & T_t(\dots, (z_{-1}, \eta_{-1}), (z_0, \eta_0), (z_1, \eta_1), \dots) \\ & = \dots, (z_{-1} + t, \eta_{-1}), (z_0 + t, \eta_0), (z_1 + t, \eta_1), \dots. \end{aligned}$$

In the case of λ arbitrary the same is true once T_t is replaced by the non-linear shift θ_t. It is defined as follows: for every $x, t \in \mathbb{R}^1$ we define $\theta_t(x) \in \mathbb{R}^1$ as the only value for which

$$\int_x^{\theta_t(x)} \lambda(x)\, dx = t.$$

Clearly,

$$\frac{d}{dt}\theta_t(x)\Big|_{t=0} = \frac{1}{\lambda(x)},$$

so the shift $\theta_t(x)$ is the same as traveling along the vector field $\frac{dx}{\lambda(x)}$ from the location x for the time duration t. We suppose that the following integrals diverge:

$$\int_{-\infty}^{0} \lambda(x)\,dx = \infty, \quad \int_{0}^{\infty} \lambda(x)\,dx = \infty, \tag{4}$$

then the measure $\lambda(x)\,dx$ is invariant under every transformation θ_t. The claims that the process $\mathcal{S}$ is invariant and ergodic under the transformation

$$\begin{aligned}&\theta_t(\ldots,(z_{-1},\eta_{-1}),(z_0,\eta_0),(z_1,\eta_1),\ldots)\\ &= \ldots,(\theta_t(z_{-1}),\eta_{-1}),(\theta_t(z_0),\eta_0),(\theta_t(z_1),\eta_1),\ldots.\end{aligned}$$

are immediate.

2.3 The exit flow

Let $\omega = (\ldots,(z_{-1},\eta_{-1}),(z_0,\eta_0),(z_1,\eta_1),\ldots) \in \Omega$. In case that for all i we have *no conflicts:*

$$z_i \geq z_{i-k} + \eta_{i-k} \text{ for all } k > 0, \tag{5}$$

we define the exit moments $E(\omega) = \{\ldots < y_{-1} < y_0 < y_1 < \ldots\}$ in the evident way by

$$y_i = z_i + \eta_i. \tag{6}$$

Otherwise we need to *resolve the conflicts.* To do so we first introduce the set $I(\omega) \subset \mathbb{Z}^1$ of all indices i, for which the relation (5) holds. Assume ω is such that the set $I(\omega)$ is double-infinite sequence $\{\ldots < i_{-1} < i_0 < i_1 < \ldots\}$. We define the sequence $R\omega = (\ldots,(Rz_{-1},\eta_{-1}),(Rz_0,\eta_0),(Rz_1,\eta_1),\ldots)$ in the following way: if $j \in I(\omega)$, then $Rz_j = z_j$. For other j we have $i_k < j < i_{k+1}$ for some k, and we put

$$Rz_j = z_{i_k} + \eta_{i_k} + \ldots + \eta_{j-1}$$

(Lindley equation). In case $R\omega$ has no conflicts, we can again use (6). Otherwise, if the set $I(R\omega)$ is again double-infinite, we can define the configuration $R^2\omega$, and so on. Assume that the configuration ω is such that

1. the configurations $R^n \omega$ are defined for all $n \geq 1$,

2. for every j the sequence $R^n z_j$ stabilizes at some finite $n = n(\omega, j)$. We denote by $\bar{R}$ the limiting transformation, which will be called the *conflict resolution operator.*

Then we define the exits $E(\omega)$ by

$$y_i = \bar{R} z_i + \eta_i.$$

Note that if the configuration $\omega \in \Omega$ violates only finitely many of the following sequence of conditions:

$$\begin{aligned} & \eta_{-1} + z_{-1} < 0, \quad \eta_{-1} + \eta_{-2} + z_{-2} < 0, ..., \\ & \sum_{i=-k}^{-1} \eta_i + z_{-k} < 0, ..., \end{aligned} \tag{7}$$

then the sequence $R^n(\omega)$ stabilizes pointwise, at every location, so the operator $\bar{R}(\omega)$ is defined. We denote by $\tilde{\Omega} \subset \Omega$ the subset of configurations thus defined, and we put $\bar{\Omega} = \cap_{-\infty<t<\infty} \theta_t \tilde{\Omega}$.

Our main assumption on the process $\mathcal{S}$ is the condition

$$\mathcal{S}\left(\bar{\Omega}\right) = \mathcal{S}\left(\cap_{-\infty<t<\infty} \theta_t \tilde{\Omega}\right) = 1. \tag{8}$$

It can be viewed as a natural generalization of the usual condition the process not to be overloaded. In what follows we will call the condition (8) the *no-overload* condition. Note that the no-overload condition implies that the probability of the server to be idle at any given time moment is positive.

The main example of the no-overloaded process $\mathcal{S}$ is obtained by imposing the following restriction on the rate of the Poisson process λ :

$$\ell \equiv \limsup_{T\to\infty} \frac{1}{T} \int_{-T}^{0} \lambda(x) \ dx < 1, \tag{9}$$

while taking $\mathcal{T}$ to be stationary and shift-ergodic, with

$$\mathbb{E}(\eta_i) \equiv 1. \tag{10}$$

The proof of that statement is the content of the Lemma 4 below. In a way, it shows that the no-overload condition is dynamically insensitive, in the sense of [BHPS].

In what follows, every point $z_i \in \omega$, such that $R^n z_i = z_i$ for all n, will be called the *point of the beginning of the cluster* of ω, or the *head of the cluster*. We will call the set

$$\mathcal{C}(z_i, \omega) = \{z_i + \eta_i, z_i + \eta_i + \eta_{i+1}, \dots, z_i + \eta_i + \eta_{i+1} \dots + \eta_{j-1}\}$$

the cluster of z_i, where $j > i$ is the index of the next after z_i point z_j of the beginning of a cluster. (In case $R\omega = \omega$, all the points z_i are heads of clusters, while each cluster contains just one point, $z_i + \eta_i$.) The segment $[z_i, z_i + \eta_i + \eta_{i+1} \dots + \eta_{j-1}]$ will be called the *support* of the cluster $\mathcal{C}(z_i, \omega)$, while the segment $[z_i + \eta_i + \eta_{i+1} \dots + \eta_{j-1}, z_j]$ – the gap between the consecutive clusters.

2.4 The Averaging Theorem

Let $b(t)$ be the rate of the process $E(\omega)$:

$$b(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\{E(\omega) \cap [t, t + \Delta t] \neq \varnothing\}.$$

The kernels $q_{\lambda,t}(x)$ are the same kernels which were used in [RS1]. They are defined as follows. Let $e(u)$ be the probability that our server is idle at the time u. (Note that the dependence of $e(u)$ on λ is only via $\{\lambda(x), x \leq u\}$.) Now define the function $c(u, x)$ as follows. Let us condition on the event that the server is idle just before time u, while at u the customer arrives. Under this condition define

$$c(u, x) = \lim_{h \searrow 0} \frac{1}{h} \Pr\left\{ \begin{array}{c} \text{the server is never idle during } [u, u+x]; \\ \text{during } [u+x, u+x+h] \text{ the server gets} \\ \text{through with some customer} \end{array} \right\}. \quad (11)$$

Then

$$q_{\lambda,t}(x) = e(t - x)\, c(t - x, x). \quad (12)$$

Theorem 1 *Let the service time process $\mathcal{T} = \{\dots, \eta_{-1}, \eta_0, \eta_1, \dots\}$ be stationary and ergodic, while the Poisson process $\mathcal{P}_\lambda$ is defined by the* ***continuous positive*** *rate function λ, such that the no-overload assumption* (8) *and condition* (4) *hold. Then for the kernels $q_{\lambda,t}$ we have*

$$b(t) = [\lambda(\cdot) * q_{\lambda,t}(\cdot)](t), \quad (13)$$

$-\infty < t < \infty$. The kernels $q_{\lambda,t}$ depend on the function λ only via restrictions $\lambda \Big|_{(-\infty,t]}$. Moreover, they are stochastic: for every t

$$\int_0^\infty q_{\lambda,t}(x)\ dx = 1, \tag{14}$$

while $q_{\lambda,t}(x) = 0$ for $x < 0$.

For the case of the process $\mathcal{T}$ to be that of independent identically distributed random variables, this theorem was proven in sections 5 and 6 of [RS1]. As in that paper, the main problem is to show the relation (14). The combinatorial counting, applied there, is valid only in the independent case.

3 Proof of the Theorem

3.1 The representation for the kernels $q_{\lambda,t}$

In this section we will obtain another representation of the kernels $q_{*,*}$, which will elucidate more clearly its stochastic nature.

Let us denote by I_x the indicator of the event that the intersection $E(\omega) \cap [x, x + \Delta x] \neq \varnothing$. Then

$$b(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_\Omega I_x(\omega)\, d\mathcal{S}(\omega).$$

By shift-invariance of $\mathcal{S}$ we also can write

$$b(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \int_\Omega \left(\int_0^1 I_x(\theta_t\omega)\, dt \right) d\mathcal{S}(\omega). \tag{15}$$

Let us fix any ω and consider all the moments $t \in [0,1]$, for which $E(\theta_t\omega) \cap [x, x + \Delta x] \neq \varnothing$. We will call them the *hitting* moments. The set of all hits will be denoted by $\tau(\omega) \subset [0,1]$. Without loss of generality we can assume that for every $t \in \tau(\omega)$ the intersection $E(\theta_t\omega) \cap [x, x + \Delta x]$ consists of just one point. (This is certainly the case, once Δx is chosen to be small enough.)

Evidently, the set $\tau(\omega) \subset [0,1]$ of hitting moments is a union of disjoint segments, $\tau(\omega) = \cup_{r=1}^{s(\omega)} D_r$. Clearly,

$$\int_0^1 I_x(\theta_t\omega)\, dt = \sum_{r=1}^{s(\omega)} l(D_r),$$

where l stays for the length of the segment. For every $t \in \tau(\omega)$ we define now the index $i(t) \in \mathbb{Z}^1$, to be the one satisfying the relation

$$\mathcal{C}\left(\theta_t z_{i(t)}, \theta_t \omega\right) \cap [x, x + \Delta x] \neq \varnothing,$$

while the index $j(t)$ is the one for which

$$y_{j(t)}(\theta_t \omega) \in [x, x + \Delta x].$$

These indices are well-defined. In words, the point $z_{i(t)} \in \omega$ is the one which becomes the head of the hitting cluster after the shift θ_t is applied, while the point $y_{j(t)}$ is just the intersection $\mathcal{C}\left(\theta_t z_{i(t)}, \theta_t \omega\right) \cap [x, x + \Delta x]$. Let us further partition the set $\tau(\omega) \subset [0, 1]$ into maximal segments of constancy of the function $i(t)$. Let us denote this partition by $\tau(\omega) = \cup_{k=1}^{u(\omega)} C_k$, while $i(k)$ will denote the (constant) value of the function $i(t)$ when $t \in C_k$. Evidently, we have

$$\int_0^1 I_x(\theta_t \omega)\, dt = \sum_{k=1}^{u(\omega)} l(C_k).$$

In general, the partition $\cup_{k=1}^{u(\omega)} C_k$ is finer than the partition $\cup_{r=1}^{s(\omega)} D_r$. However, once Δx is small enough, they are the same, provided that we know a priori that the point $y_{j(t)}(\theta_t \omega)$ moves continuously with time. In what follows we are assuming this continuity, and we postpone the proof of it till the end of the present subsection.

Let us compute the length $l(C_k)$ of the segment $C_k = [c_k, e_k] \subset [0, 1]$. At the moment $t = c_k$ the cluster $\mathcal{C}\left(\theta_t z_{i(k)}, \theta_t \omega\right)$ starts to hit the segment $[x, x + \Delta x]$, which means in particular that

$$\theta_{c_k} z_{i(k)} + \eta_{i(k)} + \eta_{i(k)+1} \ldots + \eta_{j_k} = x \tag{16}$$

for some appropriate $j_k \geq i(k)$, $j_k = j_k(\omega)$. As the time t increases from c_k to e_k, the point $\theta_t z_{i(k)} + \eta_{i(k)} + \eta_{i(k)+1} \ldots + \eta_{j_k}$ moves from the initial value x up to the terminal value $x + \Delta x$. Let us compute the time $l(C_k)$ it takes. By definition, the point $\theta_t z_{i(k)} + \eta_{i(k)} + \eta_{i(k)+1} \ldots + \eta_{j_k}$ moves with the velocity

$$\frac{1}{\lambda\left(\theta_t z_{i(k)}\right)}.$$

Therefore, we have

$$\int_{c_k}^{e_k} \frac{dt}{\lambda\left(\theta_t z_{i(k)}\right)} = \Delta x,$$

so

$$l\left(C_k\right) = e_k - c_k = \lambda\left(\theta_{\tilde{t}_k} z_{i(k)}\right)\Delta x$$

for some $\tilde{t}_k \in C_k$, due to the Mean Value Theorem. Hence, from (15) we have

$$b\left(x\right) = \int_\Omega \left(\sum_{k=1}^{\mathfrak{S}(1,\omega,x)} \lambda\left(\theta_{t_k(x)} z_{i(k)}\right)\right) d\mathcal{S}\left(\omega\right).,\tag{17}$$

where the times $t_k\left(x\right),\ 1 \leq k \leq \mathfrak{S}\left(1,\omega,x\right)$ are all the moments in $[0,1]$, at which some cluster, $\mathcal{C}\left(\theta_{t_k(x)} z_{i(k)}, \theta_{t_k(x)}\omega\right)$, contains the point x. In case $\mathfrak{S}\left(1,\omega,x\right)$ vanishes, we define the sum $\sum_{k=1}^{0}$ to be zero. Now we see that the claim (13) of our Theorem holds, with the kernel

$$q_{\lambda,x} = \int_\Omega \left(\sum_{k=1}^{\mathfrak{S}(1,\omega,x)} \delta_{\eta_{i(k)}+\eta_{i(k)+1}\cdots+\eta_{j_k}}\right) d\mathcal{S}\left(\omega\right),\tag{18}$$

see the relation (16). (Compare also with the similar relations (34) and (40) from [RS1].)

A little thought shows that indeed the r.h.s. of (18) depends strongly on $\lambda\Big|_{(-\infty,x]}$. For example, if the rate λ is small in a segment $[x_0, x]$,suitably long, then the sum of η-s in the subscript of the delta-function will typically have only one summand.

We conclude this subsection by proving the continuity statement used above.

Lemma 2 *The exit moments $y_i\left(\theta_t\omega\right)$ of the configuration $\theta_t\omega$, $i = 0, \pm1, \ldots$ – are continuous functions of t, once ω is taken from $\bar{\Omega}$.*

Proof. Let us consider only the case $i = 0$, and suppose that $y_0\left(\theta_{t=0}\omega\right) = 1$, say.

Clearly, if we would have impose the restriction that $\lambda \geq c > 0$, then our claim is immediate, since every point would move with a speed not exceeding c^{-1}. However we know only that $\lambda > 0$, so our points can have arbitrarily high speeds, and in principle it is feasible that clusters successively accelerate each other and produce an infinite speed somewhere. As we will show, that does not happen once $\omega \in \bar{\Omega}$.

Let $\mathcal{C}\left(z_{i(t)}, \theta_t\omega\right)$ be the cluster containing the exit $y_0\left(\theta_t\omega\right)$. Evidently we will be done once we know that for any T there exists a constant $K\left(T,\omega\right)$, such $i(t) \geq K(T,\omega)$ for all $t \in [0,T]$. Indeed, that means that the movement of the point $y_0(\theta_t\omega)$ is determined only by finitely many other points, while all of them have finite speeds. To see the existence of $K(T,\omega)$ we note that for any $\omega \in \bar{\Omega}$, we can write that

$$y_0(\omega) = \sup_{-\infty<k\leq 0} (z_k + \eta_k + ... + \eta_0).$$

Moreover, we know that

$$\limsup_{k\to-\infty} (z_k + \eta_k + ... + \eta_0) = -\infty. \tag{19}$$

In particular, for the location $y_0(\theta_T\omega) > y_0(\theta_0\omega) = 1$ we have for some finite $k(T)$ that

$$\theta_T z_{k(T)} + \eta_{k(T)} + ... + \eta_0 = y_0(\theta_T\omega).$$

Because of (19) we know that for some other (negative) value $K(T) < k(T)$ and for all $k \leq K(T)$

$$\theta_T z_k + \eta_k + ... + \eta_0 < 1.$$

But then we have that for all $t \leq T$

$$\theta_t z_k + \eta_k + ... + \eta_0 < 1$$

since $\theta_t z_k$ is increasing in t. That estimate means that no point $\theta_t z_k$ with $k \leq K(T)$ can be in the same cluster with the point $y_0(\theta_t\omega)$ (≥ 1) for all $t \in [0,T]$. ■

3.2 Counting of exit moments

To prove (14), it remains to establish the following

Theorem 3 *Suppose the process $\mathcal{S}$ satisfies the no-overload property* (8). *Then for every x*

$$\mathbb{E}(\mathfrak{S}(1,\omega,x)) = 1. \tag{20}$$

Proof. For $T \geq 0$ let us introduce the random variables $\mathfrak{S}(T,\omega,x)$ as the number of indices i such that $y_i(\omega) < x$ and $y_i(\theta_T\omega) > x$. For $T \leq 0$ define similarly the random variables $\mathfrak{S}'(T,\omega,x)$ to be the number of indices i such that $y_i(\omega) > x$ and $y_i(\theta_T\omega) < x$. Evidently, $\mathfrak{S}(T,\omega,x) = \mathfrak{S}'(-T,\theta_T\omega,x)$. By shift-invariance of $\mathcal{S}$, we have $\mathbb{E}\left(\mathfrak{S}'(-T,\theta_T\omega,x)\right)=\mathbb{E}\left(\mathfrak{S}'(-T,\omega,x)\right)$. So for our purposes it is enough to show that

$$\mathbb{E}\left(\mathfrak{S}\left(1,\omega,x\right)\right)\geq 1 \tag{21}$$

and

$$\mathbb{E}\left(\mathfrak{S}'(-1,\omega,x)\right)\leq 1. \tag{22}$$

If instead of termination points y-s, crossing the location x under a time-shift, we will count the number of z-points, crossing it, we obtain similarly the random variables $\mathfrak{R}(T,\omega,x)$, defined for $T\geq 0$, and $\mathfrak{R}'(T,\omega,x)$, defined for $T\leq 0$. It is immediate from our definition of the non-linear shift, that

$$\mathbb{E}\left(\mathfrak{R}\left(1,\omega,x\right)\right)=\mathbb{E}\left(\mathfrak{R}'(-1,\omega,x)\right)=1. \tag{23}$$

By ergodic theorem (see e.g. [CFS]), for $\mathcal{S}$-a.e. ω

$$\mathbb{E}\left(\mathfrak{S}\left(1,\omega,x\right)\right)=\lim_{T\to\infty}\frac{1}{T}\mathfrak{S}(T,\omega,x), \tag{24}$$

$$\mathbb{E}\left(\mathfrak{R}\left(1,\omega,x\right)\right)=\lim_{T\to\infty}\frac{1}{T}\mathfrak{R}(T,\omega,x). \tag{25}$$

Let us define the queue length $\mathfrak{Q}(\omega,x)$ as the number of indices i such that $z_i(\omega)\leq x$ and $y_i(\omega)>x$. It follows from the above definitions that for any T

$$\mathfrak{S}(T,\omega,x)\geq\mathfrak{R}(T,\omega,x)-\mathfrak{Q}(\omega,x).$$

Therefore the relations $(23)-(25)$ imply (21).

In the same way we have, that

$$\mathbb{E}\left(\mathfrak{S}'\left(-1,\omega,x\right)\right)=\lim_{T\to\infty}\frac{1}{T}\mathfrak{S}'(-T,\omega,x),$$

$$\mathbb{E}\left(\mathfrak{R}'\left(-1,\omega,x\right)\right)=\lim_{T\to\infty}\frac{1}{T}\mathfrak{R}'(-T,\omega,x),$$

and

$$\mathfrak{R}'(-T,\omega,x)\geq\mathfrak{S}'(-T,\omega,x)-\mathfrak{Q}(\omega,x),$$

which together imply (22). ■

3.3 Time integral approximation for $\mathfrak{S}$

In this section we obtain under the conditions of the Theorem 1 the infinitesimal version of the identity (20) – the identity (3). To do it, we *approximate* the function $\mathfrak{S}(T,\omega,x)$ by a time integral average,

$$\frac{1}{T}\int_0^T V_x(\theta_t(\omega))\ dt, \tag{26}$$

and we claim that the following function $V_x(\omega)$ is suitable:

i) if no cluster of ω has the point x inside its support, then $V_x(\omega)=0$,

ii) in the opposite case we have

$$z(\omega,x)+\eta_{i(\omega,x)}\ldots+\eta_{j-1}<x<z(\omega,x)+\eta_{i(\omega,x)}\ldots+\eta_j$$

for some cluster $\mathcal{C}(z(\omega,x),\omega)$ of ω and some $j=j(\omega,x)\geq i(\omega,x)$; we take

$$V_x(\omega)=\frac{1}{\lambda(z(\omega,x))\,\eta_{j(\omega,x)}}.$$

To explain the relation between the integral (26) and the number of summands in (18), let t' (resp., t'') be the moment when the above rod η_j starts (resp., ends) to cover the point x, i.e.

$$z(\theta_{t'}(\omega),x)+\eta_{i(\omega,x)}\ldots+\eta_j=x,\ \text{resp.}\ z(\theta_{t''}(\omega),x)+\eta_{i(\omega,x)}\ldots+\eta_{j-1}=x.$$

At the moment $t\in(t',t'')$ the point x moves relative to the rod η_j with velocity $(\lambda(z(\theta_t(\omega),x)))^{-1}$, hence

$$\int_{t'}^{t''}\frac{dt}{\lambda(z(\theta_t(\omega),x))}=\eta_{j(\omega,x)}.$$

Therefore

$$\int_{t'}^{t''}V_x(\theta_t(\omega))\ dt=1,$$

and so

$$\left|\int_0^T V_x(\theta_t(\omega))\ dt-\mathfrak{S}(T,\omega,x)\right|\leq 2, \tag{27}$$

where the difference is due to the influence of the rods at the ends of the interval $[0,T]$, one per end.

Therefore the expectation

$$\mathbb{E}\left(\frac{1}{T}\int_0^T V_x\left(\theta_t\left(\omega\right)\right)\ dt\right) \rightarrow 1 \text{ as } T \rightarrow \infty, \tag{28}$$

because of (20). (As we will see soon, the expectation $\mathbb{E}\left(\frac{1}{T}\int_0^T V_x\left(\theta_t\left(\omega\right)\right)\ dt\right)$ in fact equals to 1 for every $T > 0$.) On the other hand, due to the ergodic theorem,

$$\mathbb{E}\left(\frac{1}{T}\int_0^T V_x\left(\theta_t\left(\omega\right)\right)\ dt\right) = \lim_{\Upsilon\rightarrow\infty}\frac{1}{\Upsilon}\int_0^{\Upsilon}\left(\frac{1}{T}\int_0^T V_x\left(\theta_{t+\tau}\left(\omega\right)\right)\ dt\right)\ d\tau,$$

for $\mathcal{S}$-almost every ω. But, evidently, the r.h.s. limit does not depend on T, and moreover

$$\lim_{\Upsilon\rightarrow\infty}\frac{1}{\Upsilon}\int_0^{\Upsilon}\left(\frac{1}{T}\int_0^T V_x\left(\theta_{t+\tau}\left(\omega\right)\right)\ dt\right)\ d\tau = \lim_{\Upsilon\rightarrow\infty}\frac{1}{\Upsilon}\int_0^{\Upsilon} V_x\left(\theta_{\tau}\left(\omega\right)\right)\ d\tau.$$

Therefore for every $T > 0$

$$\mathbb{E}\left(\frac{1}{T}\int_0^T V_x\left(\theta_t\left(\omega\right)\right)\ dt\right) = \lim_{\Upsilon\rightarrow\infty}\frac{1}{\Upsilon}\int_0^{\Upsilon} V_x\left(\theta_{\tau}\left(\omega\right)\right)\ d\tau = 1,$$

due to (28). That proves the identity (3).

We conclude this section by presenting the proof of the

Lemma 4 *Suppose the rate λ of the Poisson process $\mathcal{P}_\lambda$ satisfies*

$$\ell \equiv \limsup_{T\rightarrow\infty}\frac{1}{T}\int_{-T}^0 \lambda\left(x\right)\ dx < 1,$$

while the process $\mathcal{T}$ is stationary and shift-ergodic, with

$$\mathbb{E}\left(\eta_i\right) \equiv 1.$$

Then the set $\bar{\Omega} \subset \Omega$ of trajectories ω, for which the conflict resolution operator $\bar{R}\left(\theta_t\omega\right)$ is defined for all t, has full measure:

$$\mathcal{S}\left(\bar{\Omega}\right) = 1.$$

This result should be compared with a classical result of Loynes, [L], see also the books [B] and [SD].

Proof. It is easy to see that the domain $\tilde{\Omega} \subset \Omega$ (see relation (8)) has full measure. Indeed, the validity of all except finitely many of the relations (7) holds $\mathcal{S}$-almost surely, due to the Strong Law of Large Numbers. This law holds here since $\mathbb{E}(\eta) = 1$, while $\ell < 1$. Moreover, the subset $\hat{\Omega} \subset \tilde{\Omega}$ of configurations ω, satisfying the relation

$$\limsup_{k\to\infty} \left(\sum_{i=-k}^{-1} \eta_i + z_{-k} \right) = -\infty, \tag{29}$$

also has full measure, for the same reason.

Next, let us show that the intersection $\bar{\Omega} = \cap_{-\infty<t<\infty} \theta_t \tilde{\Omega}$ also has full measure. Clearly, the countable intersection $\cap_{t\in\mathbb{Z}^1} \theta_t \hat{\Omega}$ has measure one. So we will be done once we show that for every T

$$\cap_{t\in\mathbb{Z}^1} \theta_t \hat{\Omega} \subset \theta_T \tilde{\Omega}.$$

This is the same as to claim that for any $\omega \in \cap_{t\in\mathbb{Z}^1} \theta_t \hat{\Omega}$ and any T we have $\theta_{-T}\omega \in \tilde{\Omega}$. This will be established once we show a stronger statement, that for any $t > 0$, any $\omega = \{(z_i, \eta_i)\} \in \hat{\Omega}$ we have $\theta_{-t}\omega \in \hat{\Omega}$. To check this inclusion we have to consider the $k \to \infty$ asymptotics of the sums

$$\sum_{i=-k}^{n(\omega,t)} \eta_i + \theta_{-t} z_{-k},$$

where $n(\omega, t)$ is the largest index i, satisfying the relation $\theta_{-t} z_i < 0$. But since $\theta_{-t} z_{-k} \leq z_{-k}$, we have evidently that

$$\limsup_{k\to\infty} \left(\sum_{i=-k}^{n(\omega,t)} \eta_i + \theta_{-t} z_{-k} \right) = -\infty$$

as well. ∎

4 Self-averaging not always holds

The example of the service discipline without self-averaging property, given in [RS1], is very simple. All the customers arriving between n *A.M.* and $n+1$

A.M. are leaving the server at $n+1$ *A.M.* sharp. Self-averaging violation is evident in this case.

However, the above discipline may not look very natural – in particular, it is not conservative. So below we present an example of conservative discipline: whenever the queue is non-empty the server is busy, and, once served, the customer leaves the server.

The example is the following. Suppose the server gets two kinds of clients: "slow" and "fast". The slow one needs time L for its service, while the fast one needs time $l \ll L$, both times are non-random. The probability that a given client turns out to be slow is $\frac{1}{2}$, say, and the sequence of service times is iid (with two values). The server has a preference for slow customers: a fast one is served only if there are no slow ones waiting in queue.

To see that the self-averaging is violated, let us take the rate function $\lambda(t)$ of the input Poisson flow to be $\Lambda \gg 1$ for $t \in [0,T]$ and zero otherwise. We are going to explain that if L, T and Λ are large enough, while l is small enough, then there exists a time moment $\tau > 0$, at which the exit flow rate $b(\tau)$ exceeds the value Λ, thus violating any possibility of self-averaging.

To see it, let us consider the random moment τ_1 of the beginning of the service of the first slow customer. Its distribution is well localized around $t=0$; in fact, it tends to δ_0 as $\Lambda \to \infty$. For every T we can find the value $L = L(T.\Lambda)$ such that the following is very likely: during the time interval $[\tau_1, T]$ the server is never idle and serves only slow customers. The time τ_2 it takes to wait untill the service of all slow clients is over is such that $\mathbb{E}(\tau_2 - \tau_1) = \frac{\Lambda}{2}LT$. At the moment τ_2 the service of waiting fast clients will start. Consider the time interval $[\tau_2, \tau_2 + \tau_3]$, during which all the fast clients are served. Its average length is $\frac{1}{2}\Lambda T l$. On the other hand, due to the Law of Large Numbers, the moment τ_2 belongs to the segment $\left[\frac{\Lambda}{2}LT(1-\varepsilon), \frac{\Lambda}{2}LT(1+\varepsilon)\right]$ with very high probability, provided that the product ΛT is large. Therefore the exit rate has to exceed the value

$$\frac{\Lambda T}{2}\frac{1}{\Lambda LT\varepsilon + \frac{1}{2}\Lambda T l} = \frac{1}{2L\varepsilon + l}$$

somewhere inside the interval $\left[\frac{\Lambda}{2}LT(1-\varepsilon), \frac{\Lambda}{2}LT(1+\varepsilon) + \frac{1}{2}\Lambda T l\right]$, which in turn exceeds Λ for l and ε small enough.

Acknowledgment. *We would like to thank our colleagues – in particular, F. Kelly, K. Khanin, Yu. Peres, S. Pirogov, O. Zeitouni – for valuable discussions and remarks, concerning this paper. A. Rybko would like to acknowledge the generous financial support of the Leverhulme Trust. He is also* *grateful to the Centre de Physique Theorique in Luminy, for the financial support during his visit in the Fall of 2004, when this work was initiated. S. Shlosman would like to acknowledge the generous financial support of the Department of Mathematics of UC Berkeley and the uplifting atmosphere of the Mathematical Sciences Research Institute during the Program "Probability, Algorithms & Statistical Physics", in the Spring, 2005, when part of that work was done.*

References

[A] Anantharam, V.: *On the rod placement theorem of Rybko and Shlosman*, Queueing Systems, 52, (2006), 185 - 188.

[CFS] Cornfeld, I. P., Fomin, S. V. and Sinai, Ya. G.: Ergodic theory. Grundlehren der mathematischen Wissenschaften [GMW] ; 245, New York, Springer-Verlag, 1982.

[B] Borovkov, A.A.: *Stochastic Processes in Queueing Theory*, Springer-Verlag, 1976.

[L] Loynes, R.M.: *The stability of a queue with nonindependent interarrival and service times*. Proc. Camb. Phil. Soc. 58 (1962), 497-520.

[BHPS] Benjamini, I., Häggström, O., Peres Yu. and Steif, J.: *Which properties of a random sequence are dynamically sensitive?* Ann. Probab. 31, no. 1 (2003), 1–34.

[RS1] Rybko, A., Shlosman, S.: *Poisson Hypothesis for Information Networks*, To appear in Moscow Mathematical Journal, 2005.

[RS2] Rybko, A., Shlosman, S.: *Poisson Hypothesis: combinatorial aspect*, Problems of Information Transmission, v. 41, # 3, p. 230 - 236, 2005.

[SD] Stoyan, D., Daley, D.: Comparison methods for queues and other stochastic models, Chichester [West Sussex] ; New York : Wiley, 1983.